AN EVALUATION OF SUBTYPING METHODS FOR IDENTIFICATION OF FECAL POLLUTION SOURCES

Alison Buchan[1], Merryl Alber[2], Mary Ann Moran[2], and Robert E. Hodson[3]

AUTHOR: [1]Graduate Student, Department of Microbiology; [2]Assistant Professor, Department of Marine Sciences; [3]Professor and Department Head, Department of Marine Sciences, The University of Georgia, Athens, Georgia 30605.

REFERENCE: *Proceedings of the 1997 Georgia Water Resources Conference*, held March 20-22, 1997, at The University of Georgia, Kathryn J. Hatcher, Editor, Institute of Ecology, The University of Georgia, Athens, Georgia 30602-2202.

Abstract. Fecal coliform concentrations are routinely monitored in waterbodies as indicators of other more potentially dangerous human fecal-borne pathogens. However, high concentrations do not necessarily indicate human waste contamination, since other warm-blooded animals also release fecal coliforms. A method to distinguish between different sources of fecal coliforms would be beneficial for identifying contamination sources and determining the level of risk to humans. In this study, two methods, one molecular and one biochemical, were evaluated to determine if host-strain specific patterns could be established between the dominant member of the fecal coliform group, *Escherichia coli*, and potential sources (i.e. human, bovine, poultry) within the watershed of Lake Sidney Lanier, Georgia. Sixteen unique isolates (3 human, 3 bovine, 4 poultry, and 6 water samples) were evaluated using these methods. The molecular method was better at discriminating between strains of *E. coli* than the biochemical, and preliminary analysis indicates that similarities between source isolates exists.

INTRODUCTION

Study Area. Lake Sidney Lanier is the primary source of drinking water for the residents of North Georgia, including most Atlantans. The Lake has been the focus of several studies of bacterial pollution (summarized in Buchan et al., submitted). Although the main body of the Lake currently has low fecal coliform counts, samples from its tributaries are routinely found to be high in fecal coliforms (often well above the state recommended 200 CFU/100ml for primary contact waters) (Hatcher et al., 1994).

The Lake Lanier drainage basin, which covers approximately 665,600 acres, is located in the heart of the poultry producing region of the state, with an estimated 274 million birds produced or housed within the basin (Smith and Sellers, 1995). This equates to the potential production of as much as 274,000 tons of poultry litter per year, all of which can potentially contribute fecal bacteria to the system (Smith and Sellers, 1995). Other potential sources of fecal coliforms include small dairy farms, waterfowl, and improperly functioning septic systems.

Background. Although fecal coliform bacteria are commonly used for determining the quality of natural waters, it is not fecal coliform *per se* that causes a problem. Rather, they are indicators of other potentially more dangerous fecal-borne pathogens such as *Salmonella* spp. and *Shigella* spp. There are, however, problems with the use of fecal coliform as an indicator of potential contamination to humans. Human waste is not the sole contributing source of coliform bacteria to a body of water. All warm-blooded animals, and even some fish, release fecal coliforms in their feces (Geldreich and Clarke, 1966). Furthermore, organisms that cause diseases in other animals are not necessarily pathogenic to people. In one study, indicator bacteria from agricultural sources were not found to be associated with human infections (Caldron et al., 1991).

Because standard fecal coliform tests group bacteria functionally, they do not distinguish between fecal coliforms from different organisms (e.g. Kelch et al., 1978, Kuhn et al., 1991). However, specific strains of *E. coli* are often associated with different host organisms, and strain-host specific interactions have been established (Faith et al., 1996). (Although, the fecal coliform group contains several types of bacteria, *E. coli* usually accounts for approximately 90% of the total individuals in any given field sample.) These findings suggest the possibility that *E. coli* strains can exhibit different characteristic patterns and thus might be traced to their source.

Related Work. Many researchers have attempted to distinguish between sources of fecal coliform in the past, with little success. These earlier methods focused on antibiotic resistance determinants (Kasper et al., 1990, Kelch and Lee 1978, and Krumperman et al. 1983), physiological differences (Kuhn et al., 1995), or fecal streptococci to fecal coliform ratios (Geldreich and Kenner, 1969). Investigators are just starting to apply molecular-based techniques to address the above problems (Herbein et al., 1996).

Overview of Method. The purpose of the present study was to compare two techniques for strain-level differentiation of fecal coliform bacteria in environmental samples in order to determine whether they are useful for discriminating between fecal bacteria from different sources (dairy cattle, poultry, and humans). First, we developed a method to distinguish between strains of *E. coli* based on slight nucleic acid sequence differences (i.e. to distinguish between these strains based on their "DNA fingerprints"). Second, we used a traditional biochemical approach to

determine whether a more conventional method would yield the same results.

The molecular method is based on differences in the 16S-23S ribosomal RNA interspacer region. The spacer region between these genes is considered a hypervariable region, and bacterial species are known to have variable DNA sequences within these regions (Jensen et al., 1993). Therefore, it has been suggested for use in identification and strain typing (Whiley et al., 1995; Jensen et al., 1993). In our method the region between the 16S and 23S genes is amplified using Polymerase Chain Reaction (PCR); with all *E. coli* strains giving two bands of approximately 480 and 540 bases (Jensen et al., 1995). We then examine these products using denaturing gradient gel electrophoresis (DGGE), which separates fragments based on their nucleic acid sequence, giving a characteristic banding pattern for each isolate.

Overview of Findings. Our results suggest that environmental *E. coli* strains can be distinguished from one another with both molecular and biochemical techniques. Moreover, preliminary analysis suggests that several strains from the same pollutant source can exhibit similar patterns. However, the biochemical results differed from those of the molecular method. We interpret these findings to suggest that differences at the molecular level are not necessarily detectable at the biochemical level.

METHODS

Study Site. Two tributaries within the Lake Lanier drainage basin, one rural (West Fork of the Little River) and one urban (Limestone Creek), were selected for our study. These sites have consistently high fecal coliform numbers (1973-present, summarized in Buchan et al., submitted), and were chosen as representative of the types of land use that are most prevalent within the drainage basin.

The West Fork of the Little River watershed encompasses a 21.5 square mile drainage basin (Hatcher et al., 1994). Land use in the watershed is divided as follows: 50% forest, 40% pasture, <7% residential, and <7% cropland (Smith and Sellers, 1995). As of 1995, this watershed contained 87 chicken houses and 5 dairy farms, with 8 animal waste lagoons (Smith and Sellers, 1995). The residential areas are classified by Smith and Sellers as low density suburban/rural with septic tanks (1995). In a section of the stream approximately 4 miles north of our sampling site (USGS monitoring station #02332830) a cattle farm is cited as a cause of large numbers of fecal coliforms in the stream (Walker, 1990).

The Limestone Creek watershed encompasses a drainage area of 18.1 square miles (Hatcher et al., 1994). Aerial photographs (1994) of the region reveal that rural livestock and poultry zones are absent from this area. Residential communities are on both sewer and septic systems. Our sampling station was located at the intersection of Limestone Parkway and Beverly Drive in Gainesville.

Isolates. *E. coli* strains were isolated from both the streams and from potential sources of fecal bacteria from samples collected on May 10, 1996. Grab samples were taken at the stream sites using sterile bottles, and *E. coli* strains were isolated using standard membrane filtration procedures on M-FC agar (Greenberg et al., 1992). Fecal samples were obtained from humans, cows, and chickens as follows: (1) a human sewage sample was taken from the influent at the Gainesville waste water treatment plant in Linwood (WPC#2) and (2) bovine and poultry fecal samples were obtained from a farm located directly upstream of the sampling site on the West Fork of the Little River. *E. coli* strains were isolated from source samples by suspending aliquots of fecal material in Tryptic Soy Broth and incubating overnight at 37°C. The broth culture was inoculated onto M-FC plates for purification and isolation of *E. coli*. All colonies expressing a blue color on M-FC agar (indicative of *E. coli*) were picked from the plates and purified using conventional methods on Tryptic Soy Agar. After isolates were purified, each was tested for their ability to utilize urease and oxidase to screen out false positives.

Strain Differentiation. In this study, we evaluated both a molecular and a biochemical method for their ability to differentiate between different strains of *E. coli*. Both techniques were applied to *E. coli* strains isolated from each source: bovine (strain designations B206, B207, B308), human (H204, H303, H205), and poultry (P107, P108, P110, P304) and from the two tributaries: West Fork (WF04, WF05, WF06) and Limestone Creek (LC02, LC03, LC06).

Molecular Method. The two primers employed for the amplification of the 16S/23S spacer region were adapted from Jensen et al. (1993). The first primer, designated G1, contains the sequence GAAGTCGTAACAAGG and is located approximately 30 to 40 bases upstream of the spacer boundary. The second primer, L1, contains the sequence CAAGGCATCCACCGT and is the most conserved sequence within the 23S gene immediately adjacent to the spacer boundary. It is located approximately 20 bases downstream from the spacer boundary. A GC clamp was attached to the 5' end of the G1 primer (CGCCCGCCGCGCCCCGCGC-CCGTCCCGCCGCCCCCGCCCCCC) so that all the products could then be distinguished from one another using denaturing gradient gel electrophoresis (DGGE) (Muzyer et al., 1993).

Amplification of the interspacer region was carried out in a reaction volume of 50µl containing: 6 µl reaction buffer (10 mM Tris-HCL [pH 8.8, 25°C] 1.5 mM $MgCl_2$, 50 mM KCL, 0.1% Triton X-100), 1 µl deoxynucleoside triphosphate (dNTP) mixture (concentration of each dNTP, 10 mM), 0.75 µl of each primer (primers G1 and L1 [concentrations 159 ng/µl and 238 ng/µl, respectively]), 3 µl

bacterial DNA in 1% Nonidet P-40 solution, and 38.5 µl deionized water. This mixture was heated to 82°C and 1 µl of Thermalase rec-Tbr (Amersco Inc.) was added, followed by 40 µl of mineral oil. PCR was carried out in a Perkin Elmer Cetus DNA Thermal Cycler 480. An initial 3 min at 94°C was followed by 25 cycles of: 1 min at 94°C; 2 min ramp to 55°C; 7 min at 55°C; 2 min ramp to 72°C; and 2 min at 72°C. The final cycle was followed by an additional 7 min at 72°C to complete any partial polymerizations (Jensen et al., 1993). All products were stored at 4°C.

The products obtained from the PCR step were then run on a denaturing gel to differentiate between slight nucleic acid differences within the amplified interspacer region. A gradient of formamide/urea (30-50%) was established and electrophoretic charge was applied in the same direction. The gel was run for 4 hours at 200 volts at 60°C. The gel was then stained with ethidium bromide and visualized on a UV transilluminator.

Biochemical Method. We employed the Biolog system (Biolog Inc.), which tests bacterial strains for the ability to utilize or oxidize 95 different carbon substrates. Each specific carbon substrate, a 150µl volume of bacterial inoculate, and a colorimetric dye was added to a 96-well microtiter plate. The plates were then incubated at 37°C for 5 hours and scored by eye as either positive, negative, or variable depending on the intensity of the colorimetric reaction.

RESULTS

Molecular. After PCR of the 16S/23S intergenic region all of the isolates exhibited the two bands (480 and 540 bp) which confirms they were indeed *E. coli*. These strains were then subjected to DGGE to separate the bands based on their nucleic acid base sequences. The results show that strains from the same source often have similar banding patterns (Figure 1). These results are encouraging in that bacteria isolated from different sources have distinct patterns. However, none of the stream isolate patterns were exactly the same as any of the source patterns. This therefore indicates that none of the stream isolates were the same strains as any of the source isolates, because this method detects minute differences at the nucleic acid level. Now that the method is optimized, the next step will be to devise DNA fingerprints for more source organisms to increase our source library. We expect this will enable us to match stream isolates with sources in the future.

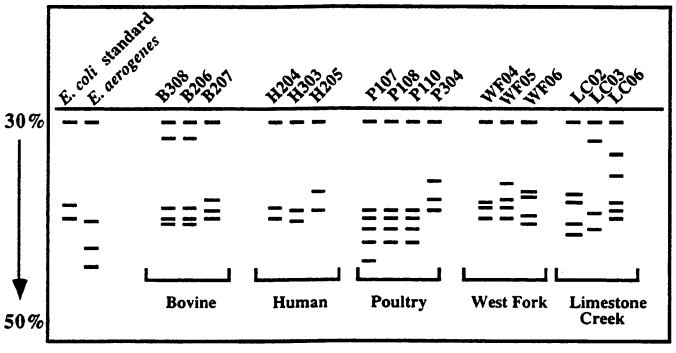

Figure 1. Line schematic of denaturing gradient gel electrophoresis. *E. coli* standard = positive control. *E. aerogenes* = negative control. A 30-50% denaturant (formamide/urea) gradient was used. Electrophoresis was run at 200 v for 4 hours in the same direction as the denaturant. The first letter of each strain name represents its source (B = bovine, H = human, P = poultry, WF= West Fork of the Little River, LC= Limestone Creek).

Biochemical. The metabolic fingerprints for the strains from each source were compared and an overall pattern was derived as follows. If all of the strains from a given pollutant source were consistent in their use of a particular substrate (e.g. either all positive or all negative), then the source pattern would be either positive or negative for that substrate. However, if strains differed in their ability to utilize a particular carbon substrate, then the source pattern was considered variable for that source. These patterns were established for all three pollutant sources. Fifty-three of the 95 carbon substrates tested were the same for the three source patterns and therefore were ignored for comparative purposes. The remaining 42 carbon substrates were then compared with the purpose of finding substrates for which: (1) all of the strains within a source group exhibited the same utilization
pattern (i.e. all positive or negative) and (2) were different between the sources. Three substrates met these criteria: succinic acid, fumaric acid, and bromo-succinic acid.

These three carbon substrates were then used to compare pollutant sources with those of stream isolates (Table 1). In contrast to the molecular results, several of the stream isolates had patterns that were similar to those of the sources (e.g. WF05 had the same pattern as the poultry source; Table 1). However, we know from the molecular results that WF05 is not in fact the same organism as any of those

Table 1. Biolog Results.

bacterial strain	succinic acid	fumaric acid	bromo-succinic acid
human	●	●	●
poultry	●	○	●
bovine	○	○	○
LC02	●	○	○
LC03	○	○	○
LC06	●	●	●
WF04	●	○	●
WF05	○	○	○
WF06	○	○	○

Human, poultry, and bovine represents source patterns obtained from various isolates from each source. The first two letters of each creek isolate indicates from where it was isolated (LC = Limestone Creek, WF = West Fork of the Little River).

● = able to utilize carbon substrate

○ = unable to utilize carbon substrate

isolated from the poultry source, because it exhibits a unique banding pattern (Figure 1). These biochemical results are the consequence of cellular physiology, which is influenced by environmental conditions such as temperature and nutrient availability. In contrast, the molecular data is based on genetic information which is not as dependent on environmental conditions.

CONCLUSIONS AND RECOMMENDATIONS

A tool to determine the origin of fecal bacteria in a sample would be useful for both assessing water quality and identifying the sources of contamination in a water body. Our results show that both molecular and biochemical approaches can be used to distinguish between strains of *E. coli*. However, biochemical assays can yield misleading results, whereas the molecular technique is less subject to fluctuations in the environment. We suggest that the molecular method shows promise in terms of differentiating between closely-related strains of *E. coli*, and will be useful for tracing pollutants to their source.

Future efforts will focus on extending the application of our results by developing molecular fingerprints for a larger number of strains from the sources indicated in this paper, as well as other potential sources (e.g. geese, horses, dogs). Once specific source patterns have been established using the molecular techniques outlined here, oligonucleotide probes based on strain-specific hypervariable regions can be constructed. This will allow for the quantification of different source of fecal coliforms in a given water sample.

ACKNOWLEDGMENTS

This work was partially funded under a USEPA Clean Lakes FY1996 Grant awarded to Robert Hodson and Merryl Alber.

REFERENCES

Buchan, A., M. Alber, M.A. Moran, and R.E. Hodson. submitted. Fecal Coliform in Lake Lanier. USEPA, Clean Lakes Program, Project Completion Report.

Caldron, R.L., E.W. Mood, and D.P. Dufour. 1991. Health Effects of Swimmers and Nonpoint Sources of Contaminated Water. *International Journal of Environmental Health Research.* 1:21-31.

Faith, N.G., J.A. Shere, R. Brosch, K.W. Arnold, S.E. Ansay, M.S. Lee, J.B. Luchansky, and C.W. Kasper. 1996. Prevalence and Clonal Nature of *Escherichia coli* 0157:h7 on Dairy Farms in Wisconsin. *Applied and Environmental Microbiology.* 62:1519-1525.

Geldreich, E.E., L.C. Best, B.A. Kenner, and D.J. Van Donsel. 1968. The Bacteriological Aspects of Stormwater Pollution. *Journal WPCF.* 40(11): 1861-1872.

Geldreich, E.E. and N.A. Clarke. 1966. Bacterial Pollution Indicators in the Intestinal Track of Freshwater Fish. *Applied and Environmental Microbiology.* 14:429-437.

Greenberg, A.E., J.J. Connors, D. Jenkins, and M.A.H. Franson., eds. 1992 *Standard Methods for the Examination of Water and Wastewater, 15th ed.* American Public Health Association, American Water Works Association, and Water Pollution Control Federation, Washington, D.C.

Hatcher, K. ed. 1994. Diagnostic/Feasibility Study of Lake Sidney Lanier, Georgia. Georgia Environmental Protection Division. USEPA, Clean Lakes Program, Project Completion Report.

Herbein, S.A., G. M Simmons Jr., and S. L. Myers. 1996. Use of Pulsed Field Gel Electrophoresis (PFGE) to Investigate Nonpoint Fecal Coliform Sources to Tidal Inlets on the Eastern Shore of the Chesapeake Bay. 6th General Meeting of the American Society for Microbiology, New Orleans, Lousiana.

Jensen, J.A., J.A. Webster, and N. Straus. 1993. Rapid Identification of Bacteria on the Basis of Polymerase Chain Reaction-Amplified Ribsomal DNA Spacer Polymorphisms. *Applied and Environmental Microbiology.* 59(4):945-952.

Kasper, C.W., J.L. Burgess, I.T. Knight and R.R. Colwell. 1990. Antibiotic Resistance Indexing of *Escherichia coli* to Identify Sources of Fecal Contamination in Water. *Canadian Journal of Microbiology.* 36:891-894.

Kelch, W.J. and J.S. Lee. 1978. Antibiotic Resistance Patterns of Gram-Negative Bacteria Isolated from Environmental Sources. *Applied and Environmental Microbiology.* 36(3):450-456.

Krumperman, P.H. 1983. Multiple Antibiotic Resistance indexing of *Escherichia coli* to Identify High-Risk Sources of Fecal Contamination of Foods. *Applied and Environmental Microbiology.* 46(1):165-170.

Kuhn, I, G.Allestam, T.A. Stenstrom, and R. Mollby. 1991. Biochemical Fingerprinting of Water Coliform Bacteria, and a New Method for Measuring Phenotypic Diversity and for Comparing Different Bacterial Populations. *Applied and Environmental Microbiology.* 57:3171-3177.

Muyzer, G., E.C. De Waal, and A.G. Uitterlinden. 1993. Profiling of Complex Microbial Populations by Denaturing Gradient Gel Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes Coding for 16SrRNA. *Applied and Environmental Microbiology.* 59:695-700.

Smith, M.C., and J. Sellers. 1995. Water Quality Impacts of Poultry Litter Management. Proceedings of the 1995 Georgia Water Resources Conference, Hatcher, K.J., ed. Carl Vinson Institute of Government, The University of Georgia, Athens, Georgia.

Walker, M.X. 1990. Water Quality Investigation of West Fork of the Little River, White and Hall Counties, Georgia. Georgia Department of Natural Resources, Environmental Proection Division. Atlanta, Georgia.

Whiley, R.A., B.A. Duke, J.M. Hardie & L.M.C. Hall. 1995. Heterogeneity among 16S-23S rRNA Intergenic Spacers of Species within the *Streptococcus milleri* group. *Microbiology.* 141:1461-1467.